On the Use of Distributional Weights in Social Cost-Benefit Analysis

Arnold C. Harberger

University of Chicago

In the cost-benefit analysis of commodity taxes or subsidies, distributional weights complicate the standard textbook "triangle analysis." Net benefits are maximized with optimal subsidies in some circumstances, optimal taxes in others. Demand and supply elasticities play important roles in determining these weighted-welfare effects of commodity taxation. I then explore how to use distributional weights (*a*) to analyze investment projects and (*b*) to determine an optimum income-tax structure. In all these applications, the use of distributional weights is shown to have very strong and (to many people) disquieting implications. A final section explores alternative "solutions" to these difficulties.

I. Introduction

In this paper I explore the use of distributional weights in applied welfare economics. Section II deals with their application to certain familiar problems of commodity taxation. Some simple results are derived which are closely related to those relating to optimum tariffs and optimal export taxes in the literature of international trade. Section III expands on the critical role which price elasticities of supply and demand play in most applications of distributional weights in cost-benefit work. Three cases are set out which in a sense dramatize this role, and an effort is made to help the reader reach an intuitive understanding of it.

Section IV explores how the use of distributional weights affects the cost-benefit analysis of investment projects, while Section V shows how

This paper was presented at a conference on Research in Taxation sponsored by the National Science Foundation and the National Bureau of Economic Research, Stanford, California, January 1976.

[*Journal of Political Economy*, 1978, vol. 86, no. 2, pt. 2]

these weights come into play in some rudimentary problems of income taxation.

Section VI tries to distill the (somewhat disquieting) messages that emerge from the preceding sections into a basic dilemma with which the use of distributional weights almost inevitably confronts us. Finally, in Section VII, three possible alternative "solutions" to this dilemma are presented and explored.

II. Welfare Analysis of Commodity Taxes

In the classical analysis of the efficiency costs of a commodity excise tax, the measurement can be approached in two distinct ways. In figure 1*a* the problem is viewed from the angle of the quantity axis. The reduction of consumption of $X_1^0\ X_1^1$ units entails a loss of output which is valued by consumers as the (vertically shaded) area under the demand curve between X_1^0 and X_1^1. The resources released have a value, at their marginal supply price to this use, equal to the (horizontally shaded) area under the supply curve. This also measures the value that the owners of these resources place on the alternative use to which the resources will be shifted in the presence of the tax, which in turn is equal to the social benefit of having them in those uses as long as either (*a*) there are no distortions other than the tax being analyzed, or (*b*) the distortions which exist in the rest of the economy generate "external" costs and benefits whose net effect is zero (or small in the context of the problem under study).[1] The net effect is the standard triangle of excess burden, *ABE*.

The second way of looking at the problem is from the angle of the price axis, as shown in figure 1*b*. Here the loss of consumer surplus ($C-$) generated by the tax is given by the vertically shaded area $ABP_1^dP_1^o$, while the loss of producer surplus ($F-$) is the horizontally shaded area $AEP_1^sP_1^o$. Since the gain ($G+$) to the government (other taxpayers, transfer recipients) is given by $BEP_1^sP_1^d$, there is a net loss of uncompensated consumer surplus equal to *ABH* and a net loss of uncompensated producer surplus equal to *AEH*. Together these give exactly the triangle *ABE* derived by the alternate method in figure 1*a*.

[1] The general expression for these "external effects" is

$$\int_{T_1=0}^{T_1^*} \sum_{i=2}^{N} D_i \frac{\partial X_i}{\partial T_1} dT_1,$$

where D_i measures the excess of marginal social benefit over marginal social cost in activity i. If D_i is constant for the problem in question (i.e., does not vary as a consequence of changes in T_1 over the range from zero to T_1^*), this expression reduces to $\sum_{i=2}^{N} D_i \Delta X_i$. See Harberger (1971).

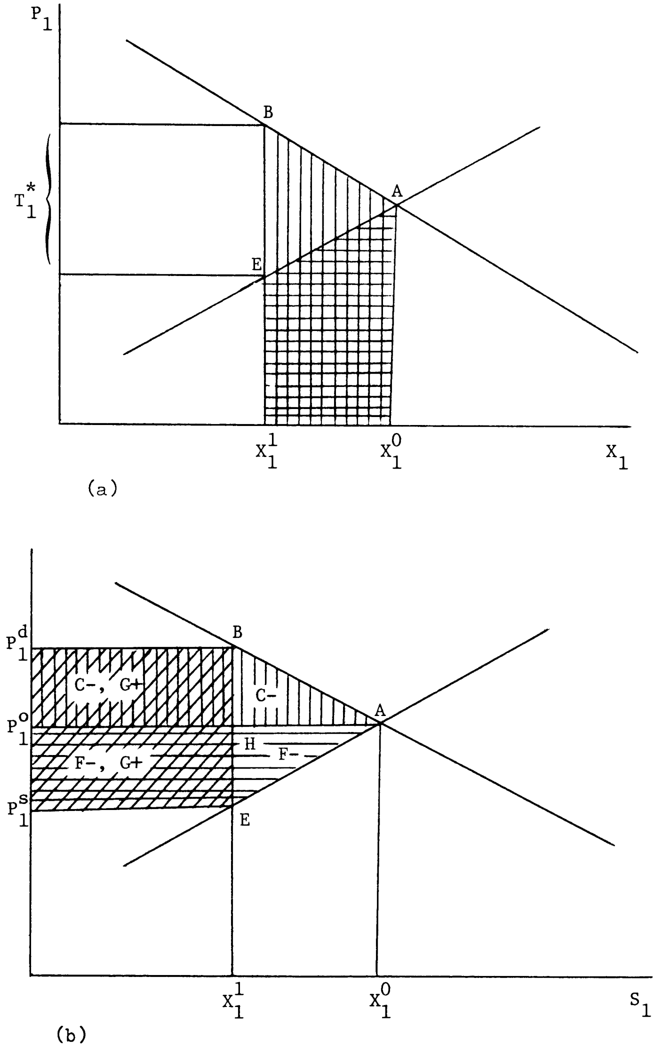

FIG. 1

Both of these methods make use of the postulate that the monetary value of benefits (+) and costs (–) is summed algebraically across the relevant individuals and groups. The method of figure 1*a* uses it directly in netting out the area $EAX_1^0X_1^1$ and indirectly in failing even to mention the implicit transfers shown as ($C-$, $G+$) and ($F-$, $G+$) which constitute the way in which the postulate is reflected in the method of figure 1*b*.

In considering how the traditional evaluation of welfare costs changes when distributional weights are applied, we adopt the convention that the standard weight (ϕ) is unity and that this weight applies to changes in the welfare of the average citizen as well as of the government. When a particular group (i) is to be preferred, as against the average citizen, it is assigned a welfare weight ϕ_i greater than one, and when it is to be "dispreferred" (not to use more pejorative terms like "punished" or "discriminated against"), its ϕ_i will be less than one.

A little reflection will show that, when distributional weights are applied, the method of figure 1a cannot be used, while that of figure 1b must be adapted to accommodate the change of assumptions. The reason is that the only "base" to which a distributional weight can reasonably be applied, in an analysis of a group's welfare, is the change in economic surplus experienced by that group.

The way in which distributional weights work, when applied to the classical tax problem, can be illustrated by considering the case in which the demanders of X_1 constitute a "dispreferred" group, while the suppliers are "average citizens." In this case, the analysis surrounding figure 1b would remain the same for the shaded areas lying below P_1^oA but would be altered for the shaded areas above that line. Now the loss to consumers would be given smaller weight than the gain to the government. With distributional weight of ϕ_c for consumers and of unity for the government and for suppliers, the benefit-cost picture now consists of a loss of the full triangle AEH, a loss equal to a fraction ϕ_c of the triangle ABH, and a gain equal to $(1 - \phi_c)$ times the transfer rectangle $HBP_1^dP_1^o$. Needless to say, under these circumstances the net result can be a gain.

A natural extension of this line of reasoning is to find the optimum tax on X_1, that is, the level of tax, $\hat{T}_1$, which produces the largest social gain, assuming the effects of other distortions to be absent or negligible. This case is illustrated in figure 2a. For a small change in the tax dT_1, the following changes in costs and benefits are generated.

1. The lower "flat" rectangle labeled $(F-, G+)$. For the present problem this generates no change in social benefit because the weights attaching to the owners of the relevant productive factors (F) and to the government (G) are equal.

2. The upper "flat" rectangle labeled $(C-, G+)$. This generates a net social gain equal to $(1 - \phi_c)X_1dP_1^d$, because the distributional weight attaching to consumers is less than that applied to changes in revenue to the government.

3. The two "tall" rectangles labeled $(G-)$. These generate net costs equal to T_1dX_1. The lower of these rectangles is part of the loss to supplying factors, and the upper is part of the loss to demanders, under both T_1^0 and T_1^1. Hence these areas do not contribute to the changes in welfare of these

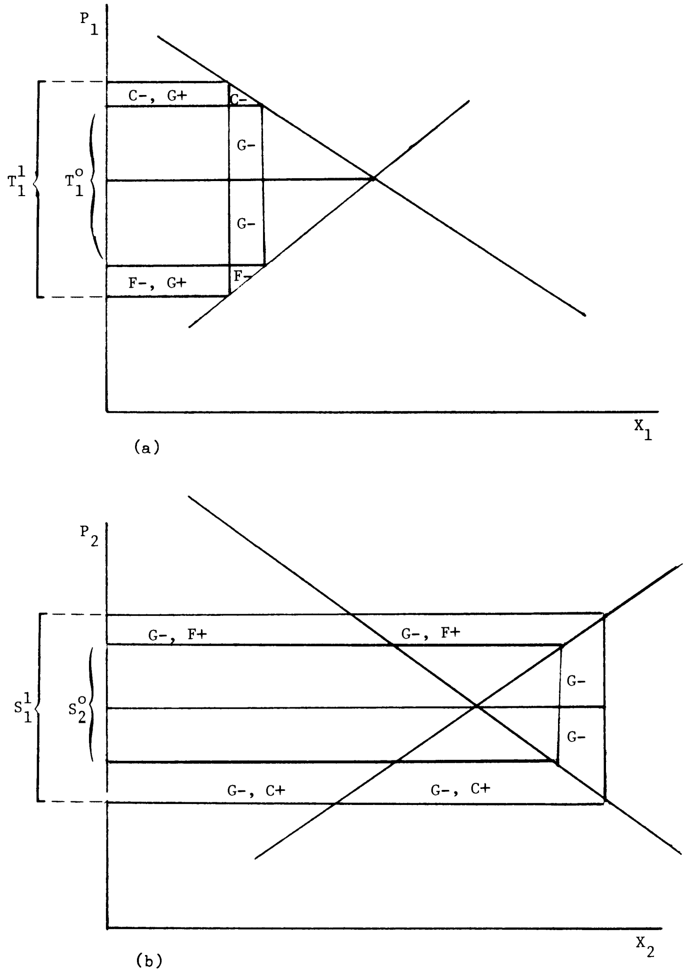

FIG. 2

private groups, as between the two tax situations. But they do constitute negative components of the change in the command over resources of the government; hence their label $(G-)$.[2] The quantitative measure of this contribution to the change in social welfare generated by a small change in the tax is $T_1 dX_1$ (with negative sign because $dX_1 < 0$).

[2]Another way to see why these areas are included as cost items is to consider that they pass from being parts of the "transfer" rectangles [$(C-, G+)$, $(F-, G+)$] to being parts of the "deadweight loss" triangles [$(C-)$, $(F-)$] of a diagram like fig. 1*b*. In the process, the only group whose net position changes, as far as these particular areas are concerned, is the government.

4. The two small triangles labeled $(C-)$ and $(F-)$. These are additions to the standard "excess burden" triangle of the traditional analysis. When distributional weights are used, those corresponding to consumers and factor suppliers should be applied, in principle, in order to arrive at the contributions of these triangles to the overall measure of social welfare. However, when, as here, the objective is to find an optimum position, the triangles in question are left out of account because they are of a second order of smalls.

The net change in social welfare stemming from a small change, dT_1, in the tax on X_1 is therefore

$$\Delta W = (1 - \phi_c) X_1 \, dP_1^d + T_1 \, dX_1. \tag{1}$$

Setting this equal to zero and defining $\hat{\tau}_1 = \hat{T}_1 / P_1^d$ and $\eta_1 = (P_1^d / X_1)(dX_1 / dP_1^d)$, we have

$$\Delta W = (1 - \phi_c) X_1 \, dP_1^d + \hat{T}_1 \, dX_1 = 0 = [(1 - \phi_c)/\eta_1] + \hat{\tau}_1. \tag{2}$$

The optimum rate of tax is therefore $\hat{\tau}_1 = -(1 - \phi_c)/\eta_1$, that is, one minus the welfare weight of the consuming group, divided by the absolute value of their elasticity of demand for the product in question.

Figures 1*b* and 2*a* can also be used to represent the case where the demanders are ordinary citizens and the suppliers are the "dispreferred" group. Now, in figure 1*b*, it is the rectangle labeled $(C-, G+)$ which cancels out, while the rectangle $(F-, G+)$ contributes a positive benefit equal to the fraction $(1 - \phi_f)$ of its area. The loss components are the entire triangle labeled $C-$ and the fraction ϕ_f of that labeled $F-$. In finding the optimum tax in this case, the relevant equation is

$$\Delta W = -(1 - \phi_f) X_1 \, dP_1^s + \hat{T}_1 \, dX_1 = -[(1 - \phi_f)/\varepsilon_1] + \hat{\tau}_1 = 0.^3 \tag{3}$$

Readers with a background in international trade theory may at this point have a sense of déjà vu. For the twin problems of (*a*) the optimal export tax and (*b*) the optimal import tariff on a single commodity, as treated in that theory, have closely similar solutions. The optimal export tax, expressed as a fraction of the gross-of-tax price, is $\hat{\tau}_i = -1/\eta_i$, where η_i is the elasticity of foreign demand for this country's exports of good *i*. Similarly, the optimal import tax, expressed as a fraction of the net-of-tax price, is $\hat{\tau}_j = 1/\varepsilon_j$, where ε_j is the elasticity of the world excess-supply function of good *j* facing this country's importers. These taxes exploit, respectively, the monopoly power that this country has in the world market for its exports of good *i* and the monopsony power that it has vis-à-vis the world-market supply of its imports of good *j*. It is well

[3] Here we define ε_1 as the elasticity of supply $(P_1^s/X_1)(dX_1/dP_1^s)$ and $\hat{\tau}$ as $\hat{T}/P_1^s$.

known, too, that the "optimum tariff" and "optimal export tax" are indeed optimal only from the standpoint of the country in question. From a world-welfare point of view they constitute distortions which taken by themselves would work only to reduce economic efficiency and welfare.

Our results coincide exactly with those of international trade theory. When the welfare weight of the consumers of a commodity becomes zero, our optimum tax, $\hat{\tau}_i = -(1 - \phi_i)/\eta_i$, on that commodity reduces to the optimal export tax of trade theory. Likewise, when the welfare weight of the suppliers of a commodity becomes zero, our expression for the optimum tax, $\hat{\tau}_j = (1 - \phi_f)/\varepsilon_j$, reduces to the formula for the optimal import tariff. It is quite clear, then, that the grand tradition of international trade theory already has dabbled with distributional weights, at least in a rudimentary sort of way, by assigning, for certain problems, a welfare weight of zero to foreigners. It is also gratifying to know that, for similar underlying assumptions, its answers and those of our own analysis are identical.

The simplicity, and in this sense elegance, of the optimum-tax solution extends to the problem of optimum subsidies, a subject left untouched by the classical literature on trade. In this case we are dealing with situations in which the suppliers and/or demanders of a commodity constitute a "preferred" rather than "dispreferred" group. Figure 2*b* is an illustration that is relevant to either case. When the consumers of the commodity in question are a "preferred" group ($\phi_c > 1$), while the suppliers are just "ordinary citizens" ($\phi_f = 1$), the areas marked ($G-$, $C+$) generate positive contributions to social welfare as a consequence of a small increase dS_2 in the subsidy on X_2. The areas marked ($G-$, $F+$), on the other hand, reflect pure transfers between groups of equal weight and hence generate no net benefit or cost. Finally, the areas marked ($G-$) represent increments in costs to the government, uncompensated by benefits to either suppliers or demanders. The equation for the change in welfare, which in turn defines the optimal subsidy $\hat{S}_2$, is

$$\Delta W = -(\phi_c - 1)X_2\, dP_2^d - \hat{S}_2\, dX_2 = 0 = (\phi_c - 1)/\eta_2 + \hat{\sigma}_2 . \quad (4)$$

Here $\hat{\sigma}_2$ is the optimal percentage rate of subsidy $[(\hat{S}_2/P_2^d) = -(\phi_c - 1)/\eta_2]$, and η_2 is the elasticity of demand for X_2.

When the suppliers of factor services are the "preferred" group ($\phi_f > 1$) and demanders have a unitary distributional weight, the corresponding relation is

$$\Delta W = (\phi_f - 1)X_2\, dP_2^s - \hat{S}_2\, dX_2 = 0 = [(\phi_f - 1)/\varepsilon_2] - \hat{\sigma}_2 . \quad (5)$$

Here ε_2 is the elasticity of supply of X_2 and the optimal percentage rate of subsidy is $\hat{\sigma}_2 = (\phi_f - 1)/\varepsilon_2$.

Not surprisingly, just as the optimum rate of commodity tax turns out to be a partial exploitation of monopoly or monopsony power vis-à-vis a "dispreferred" group, the optimum rate of subsidy turns out to be a manifestation of a sort of monopsonistic or monopolistic preference toward a "preferred" group. That is to say, whereas the exploitation of a monopoly position entails raising the price to demanders by a percentage equal to the reciprocal ($1/|\eta_i|$) of the demand elasticity, a situation in which $1 > \phi_c > 0$ calls for raising price to demanders by a fraction of the monopoly percentage, the fraction being $(1 - \phi_c)$. When $\phi_c > 1$, however, optimization calls for setting a price that is below marginal production cost—a most unmonopolistic behavior. Yet the traces of the monopoly situation are present, in the sense that the percentage by which price should fall short of marginal cost here once again will vary inversely with the elasticity of demand. When demand is highly elastic, the optimum subsidy will be negligible, even though the consuming group is quite strongly preferred (say $\phi_c = 2$ or 3 or 4). On the other hand, the optimum percentage subsidy can be quite high, even when the demander group is only modestly preferred (say $\phi_c = 1.2$, 1.3, or 1.4), as long as the elasticity of demand is low enough. The same principle applies, mutatis mutandis, in cases where suppliers are the preferred group, the relevant elasticity in this case being that of supply.

I consider the role played by supply and demand elasticities to be a key feature of the use of distributional weights in the analysis of taxes and subsidies as well as in other branches of applied welfare economics. Because of its importance, this role is dealt with in greater detail in the next section.

III. The Interaction of Weights and Elasticities

The bulk of this section (parts A, B, and C) shows how distributional weights should be applied (if they are used at all) in the analysis of three simple and familiar problems of policy economics. In part D, some general lessons are extracted from the earlier exercises.

A. Consider the familiar public-finance problem of designing a progressive excise-tax structure, say for a less developed country whose income-tax administration is not sufficiently strong to warrant its carrying the entire progressive component of the tax system. The common approach to this problem is to array commodities according to their cross-sectional income elasticities, taxing at the highest rates those whose consumption is most highly concentrated in the upper income groups and at progressively lower rates those with lower cross-sectional income elasticities. In terms of the present analysis, this procedure corresponds to attaching distributional weights to the various income brackets (with the weights being less than unity in the upper range of the income distribution and declining as income rises) and taxing commodities at rates which vary inversely with the average distributional weight of those who consume them.

The results of the previous section suggest, however, that this solution is too simplistic. The proper solution would take price elasticities of demand into account as well. If the commodities in question were independent on the demand side, the optimum pattern of taxation would be $\hat{\tau}_i = -(1-\bar{\phi}_i)/\eta_i$, where $\bar{\phi}_i$ is the average distributional weight of the consumers of X_i. Such a pattern could easily diverge from the more conventional one, owing to variations in η_i across commodities. In particular, it is easy to imagine that a good like refrigerators (moderate $\bar{\phi}_i$, low η_i) might end up with a significantly higher optimal tax than, say, luxury cars (low $\bar{\phi}_i$, high η_i).

B. Consider now a hypothetical case in which the government adopts a policy of subsidizing certain agricultural products produced, in the main, by poor farmers. Let us assume, for the sake of the argument, that to be admissible for such a subsidy 75 percent of the production of a commodity would have to be carried out by farmers classified as being below the poverty level, according to certain indicators of socioeconomic status. For our purposes, we can take this to mean that for each commodity meeting the standard (say, barley, corn, and rice) the average distributional weight ($\bar{\phi}_j$) of the supplying farmers would be the same—perhaps something like 1.5.

The casual or conventional response in this case would be to set an equal rate of subsidy on the three commodities in question. But that would be as wrong as was a uniform tax in the problem of case A. The correct answer, given that it has been decided to employ distributional weights, is that the subsidies should vary in inverse proportion to the elasticities of supply of the three goods. If $\bar{\phi}_j$ is indeed 1.5 for all three, and if the elasticity of supply is 1.0 for barley, 1.5 for corn, and 2.0 for rice, then the optimal subsidy pattern would be a 50 percent subsidy on barley, one of 33⅓ percent for corn, and one of 25 percent for rice. It is easy, of course, to imagine real-world examples in which the variation among optimal subsidy rates, owing to differential supply elasticities, would be much wider than this.

C. Here we consider a case from the expenditure rather than the tax side of public finance. Assume that in the evaluation of public-sector projects and programs it has been decided to employ distributional weights and that weights in excess of unity are applied to unskilled and low-skilled labor. Figure 3 illustrates a case where a public program employs L_1L_2 of such workers at the market wage rate. The wages bill of the program is equal to L_1L_2BC, the increment of employment is L_0L_2; and the amount earned by this increment is L_0L_2BD; the overall increase in the earnings of this class of workers is $W_0AL_0L_2BW_1$. Yet none of these represents the benefit to which the distributional weight premium $(\bar{\phi}-1)$ should be applied. The relevant base for that premium is the gain in the surplus of the affected group of workers—the diagonally shaded area in figure 3. Note that only a small part of this benefit (the triangle ABD) accrues to the newly employed workers and a not much larger part (the

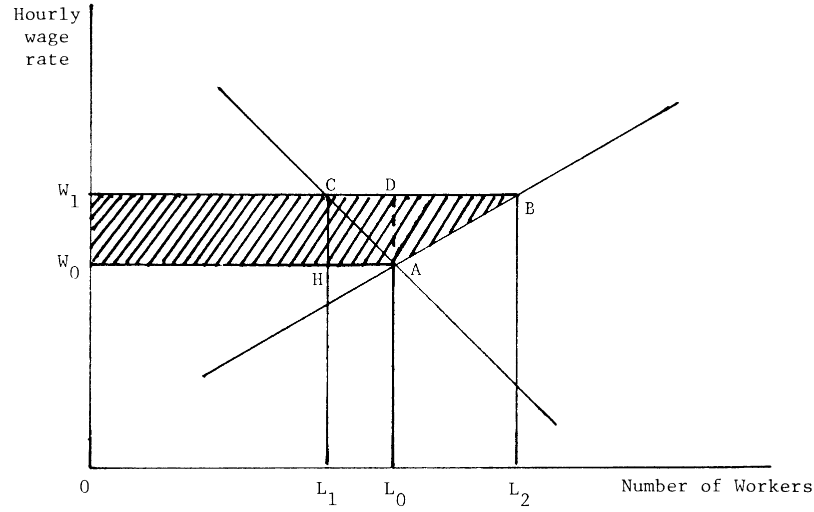

FIG. 3

trapezoid $ABCH$) to the whole labor force (L_1L_2) of the project in question.[4] The bulk of the distributional benefit (W_0W_1CH) goes to workers who would in any event have been employed elsewhere and who, even in the presence of the project, are not employed by it.

The size of the true gain in surplus of the affected group of workers is related to the wages bill of the project (L_1L_2BC) by certain elasticity relationships. First, note that $LdW = (1/\varepsilon)WdL$ for movements along the supply curve; therefore, the area ABW_1W_0 is (approximately) equal to $(1/\varepsilon)$ times L_0L_2BD. This latter rectangle in turn is equal to the fraction $\varepsilon/(\varepsilon - \eta)$ of the total wages bill of the project, L_1L_2BC. Hence the area that is relevant for analysis of distributional benefits of such a "direct-hire" project is equal to $1/(\varepsilon - \eta)$ times the actual outlays of the project for the kind of labor in question.

In this case both the elasticity of supply and that of demand (by other demanders than the project itself) enter into the picture, and they enter in a symmetrical way. Thus it is not at all true that it is a matter of indifference to society whether a given amount will be spent to hire unskilled or low-skilled labor on one project or another. When distributional weights are employed, the benefit that is attributable to a project because of their use could be vastly different, even for the same wages bill, depending on the elasticities of supply and demand in the particular (occupational, regional, etc.) labor market in question.

[4]The ratio of $ABCH$ to the total shaded area is less than the fraction of the total labor force in question that is employed by this project. For most practical cases this fraction would be tiny (rarely as high as 5 or 10 percent).

D. In each of the three cases just discussed, the distributional benefits of the policy under review depended on elasticities of supply and/or demand. In particular, such benefits were shown to be zero if the relevant elasticity were infinite. This is an old truth, but one which nonetheless many people find hard to swallow. If eggs are in infinitely elastic supply, a subsidy to their production will do poultry farmers no good; if construction workers are in infinitely elastic supply, a project which increases demand (at the supply price) for their services will do them no good; likewise, demanders will not be hurt by a tax on an item for which very close substitutes are available to which users may costlessly turn as the supply of the taxed item dwindles. All these situations reflect the fact that supply and demand prices reflect indifference, on the part of economic agents, between marginal units of the good or service in question and marginal units of the most immediate and economically interesting alternatives. When the elasticities are infinite, this property applies to inframarginal and extramarginal units as well, over whatever range is covered by the infinite elasticity. When the supply of eggs is infinitely elastic, a subsidy benefits consumers, not producers. When more workers are hired by a construction project, they will be benefited to the extent that the wage they are actually paid is above their supply price, and inframarginal workers will be benefited to the extent that the wage is further bid up above their supply price, but none of these groups will have its welfare improved if the new workers receive just barely their supply price and if the wage paid to the "old" workers does not change. Where a taxed item is in infinitely elastic demand, it is suppliers, not demanders, who are hurt by the tax.

Two examples drawn from the situations of some less developed countries may serve to reinforce the point. The first concerns the capacity of these countries to implement heavy or strongly progressive taxation of the income from capital. In this case the world capital markets have an infinitely elastic demand for the funds of LDC investors. There being no effective way for the LDCs to tax the income generated by such investments made by their nationals, the rate obtainable on these investments provides a floor below which the governments cannot readily drive the net-of-tax return on investments made at home. Thus heavier corporation income taxes in Panama or Guatemala simply mean less capital invested in the corporate sectors there and very likely more capital invested by nationals of those two countries on Wall Street or in the Eurodollar market. The net yield to Panamanian or Guatemalan investors on their investments at home would probably not be affected for long by a rise in the corporation tax rate. Instead, the gross yield would likely rise to reflect the tax, which would thus end up impinging more on the welfare of local consumers than on that of capitalists.

The second case concerns policies designed to deal with the problems of rural underemployment and poverty in certain LDCs. In some places the opinion is widespread that such underemployment is manifested neither in open unemployment nor simply in a low equilibrium level of real wages. Instead, this line of thought runs, it is reflected in a highly elastic supply of unskilled labor, unwilling to work below the going "conventional" wage but willing to work at it. Accepting these assertions as facts, the welfare-economics conclusion would be that programs to augment the demand for such labor would produce no distributional benefit and could be justified only if they could "bear" a full costing of the labor they employed.

Programs designed to fight poverty by augmenting the demand for labor really come into their own when the facts are quite different—that is, when the conventional market supply and demand curves are quite inelastic. This is a direct application of the analysis surrounding figure 3. The true power of such programs is measured by their capacity, per unit of outlay, to bid up the general level of wages for the given type of labor in the labor market in question. And, as was shown above, the bulk of distributional gains they produce accrue to workers having no direct involvement with the project.

IV. Distributional Weights and the Social Evaluation of Investment Projects

This subject has already been touched upon (Sec. III.C above) in connection with measuring the distributional benefit of a project that increases the demand for labor of a type that is judged to merit favorable treatment (i.e., with a distributional weight $\phi_i > 1$). In that case it was seen that the distributional effect was mainly external to the project itself, hinging upon whether the market wage for workers of the type in question was or was not raised by the presence of the project. Where in fact the market wage was raised, it was seen that the benefits would in most instances accrue mainly to workers employed in jobs having no direct connection with the project. The distributional benefits were in this sense indirect and external to the project in question. In this section we examine situations in which distributional weights impinge more on the direct benefits (Sec. IV.A) or costs (Sec. IV.B) of a project.

A. For the benefit side, consider a class of projects (public roads, parks, etc.) whose benefits accrue mainly "in kind," and for which the beneficiaries do not have to pay a charge or fee. Assume, furthermore, that projects of this class can be replicated at constant cost and that there are no interrelationships in demand among the separate projects (in the sense that the benefits and costs of any one project are independent of whether or not any of the others is undertaken).[5]

[5] This set of assumptions is not essential, but it is useful because it serves the didactic purpose of enabling me to focus directly on the point I am trying to make.

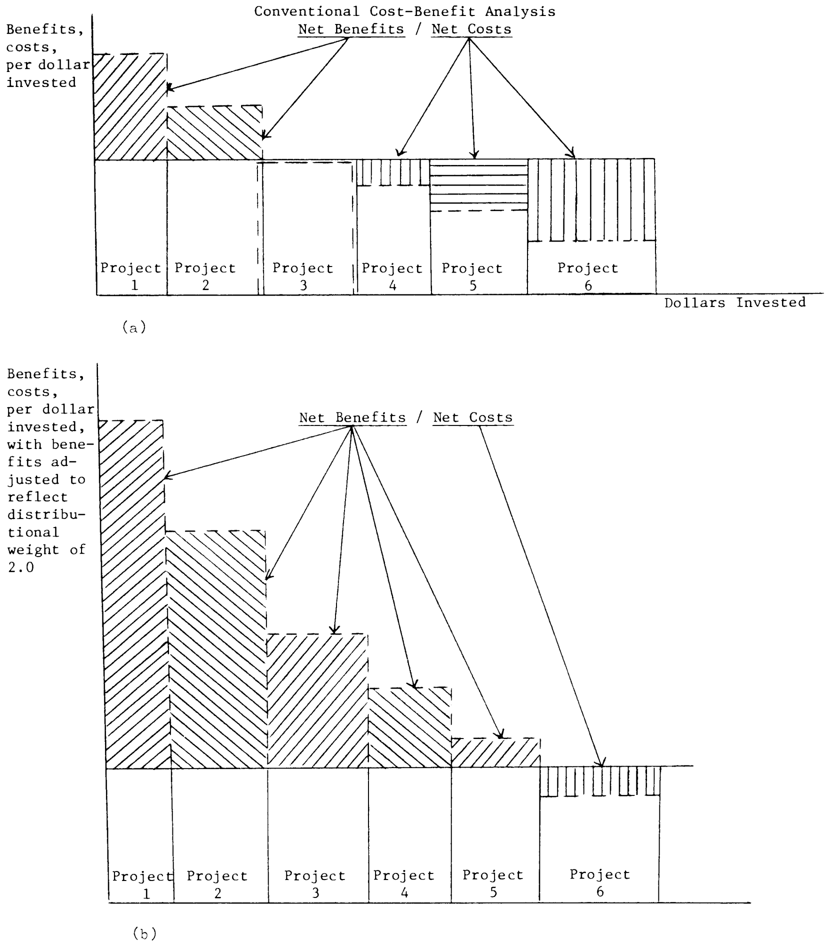

FIG. 4.—*a*, Conventional cost-benefit analysis. *b*, Same projects, with distributional weight of 2.0 for beneficiaries.

These assumptions are reflected in figure 4. Here a series of investment projects is arrayed along the horizontal axis. It is assumed that no complications (distributional or other) arise on the cost side, so that the standard methodology of social cost-benefit analysis can be applied. Each project's cost is then measured both by the width of its bar and by the rectangle enclosed by the heavy line. Each project's benefits are measured by the rectangle enclosed by the dotted line. Positive net benefits are represented by the areas shaded diagonally; negative net benefits by the areas shaded vertically or horizontally.

It is easy to see in figure 4*a* that only projects 1 and 2 yield a social "surplus" according to standard criteria; on this basis project 3 is at the margin, and projects 4, 5, and 6 are not worthwhile.

Once a distributional weight of 2 is applied to the beneficiaries, however, all but project 6 pass the test of social profitability. The critical point here is that now projects pass the test as long as their unweighted benefits amount to more than half their costs. That is to say, the new test accepts a "waste" of up to half the resources devoted to a project as a legitimate "price" to pay for the income transfers that the project carries with it.

It is important that the reader appreciate the sense in which waste is involved. If, for example, the beneficiaries of each project simply were "average citizens," with distributional weights ϕ_i equal to one, then figure 4*a* would apply, and the acceptance of projects 4 and 5 would genuinely imply the social costs (shaded areas) depicted there for those projects.

Alternatively, if it were known that the projects in question would be financed by additional taxes on people whose distributional weight was the same (2.0) as that of the beneficiaries, then projects 4 and 5 would again be rejected by the social-evaluation criterion. This time their social costs would be twice the areas shown for these projects in figure 4*a*—twice because in this case both the benefits and the costs of each project would be multiplied by the distributional weight of two.

Thus we see that the only reason why projects 4 and 5 turn out to be acceptable in figure 4*b* is because the beneficiaries have a higher distributional weight than those who are assumed to pay the costs of the project. Society ends up, by the criterion of distributional weights, paying a price in terms of efficiency for each incremental distributional benefit it gets.

This seemingly modest point is of extreme importance, for it introduces a whole new element into social-project evaluation, making the evaluation of the project itself depend in a critical way upon the manner in which the project is financed. One of the notable virtues of the traditional methodology—the fact that its measure of the social profitability of a project is largely independent of how the project is financed[6]—is therefore lost.

B. Turning now to cases in which distributional benefits come into play with respect to the direct costs of a project, let us consider once again a situation in which the workers hired on a project have a distributional weight greater than one. But here, as distinct from what was done in Section III.C, let it be true that the projects in question pay a higher-than-

[6]The method of financing enters to a degree even in the traditional analysis. For example, the benefits and costs of a given road will differ if a toll is charged or not. Also, a low-interest-rate loan from foreigners which is "tied" to a particular project may under certain circumstances give rise to an external benefit attributable to that project. But instances of these types are likely to be infrequent and their impact on the measured social merit of a project small in comparison with what happens under the generalized use of distributional weights.

market wage. Under these conditions, it is likely that the marginal worker employed by the projects will perceive some economic rent, giving rise to a possible distributional benefit (if the relevant ϕ_i is > 1) or cost ($\phi_i < 1$).

To isolate the effect in question from that treated in Section III.C, let us assume that we are dealing with a type of labor that is in infinitely elastic supply (over the relevant range). This rules out a new project's causing the market wage to be higher than it otherwise would be, and at the same time establishes unambiguously the private opportunity cost of any workers hired by the project. For the sake of simplicity, assume that the only costs of the projects considered are for labor of the type in question and that such labor is paid (for working on the projects) 1.5 times the wage it could get elsewhere. The above assumptions are reflected in figure 5*a*, where the conventional procedures of cost-benefit analysis are applied. Barring complications other than those stated, the social calculus of costs would differ from the strictly monetary one in that labor would be valued at its alternative earnings of $1.00 rather than the wage actually paid. This is equivalent to assigning to each worker employed a benefit equal to the economic rent ($0.50) he receives. While only project 7 would pass the strictly monetary cost-benefit test (assuming that project benefits came in the form of cash), both 7 and 8 would pass the conventional social cost-benefit evaluation. Projects 9, 10, 11, and 12 would fail both tests.

Figure 5*b* depicts the same projects modified to incorporate a distributional weight of 2.0 for workers. This weight must be applied to the amount by which the earnings of labor (taken in all cases to be $1.50) exceed the opportunity cost as seen by the workers in question. This economic rent is still $0.50, as in the conventional analysis. But here a social valuation of $1.00 is placed on the economic rent of workers, in the light of their distributional weight of 2.0. Thus, in the final reckoning, the labor which is being paid $1.50 has a net social cost of only $0.50. This shifts projects 9, 10, and 11 into the acceptable category, leaving only project 12 with adjusted costs greater than benefits.

Clearly, the application of distributional weights makes the difference between figures 5*a* and 5*b*, but its full meaning may not yet be clear. To explore a bit further, assume now that the wage paid by the projects was $1.00 rather than $1.50. Now the laborers would earn no economic rent, hence they would have no benefit to which to apply the distributional weight. Their social opportunity cost would equal their alternative earnings of $1.00 (regardless of whether a distributional weight were assigned or not, and regardless of the magnitude of whatever weight might be assigned), and only projects 7 and 8 would pass muster.

The argument of the preceding paragraph points up the most troublesome aspect of distributional weights as they impinge on project costs. So long as some costs are for factors whose suppliers have distributional weights greater than one, it is possible to raise the net benefits of a project

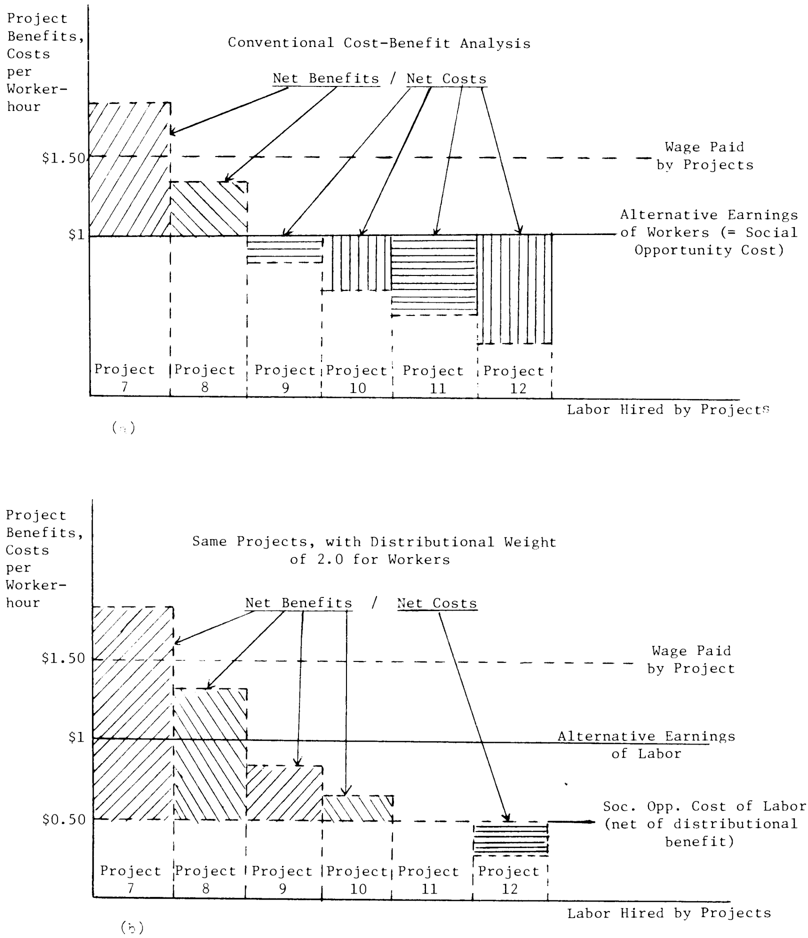

FIG. 5.—*a*, Conventional cost-benefit analysis. *b*, Same projects, with distributional weight of 2.0 for workers.

simply by raising the wage (or other price) paid to such factors. Moving from a wage of $1.00 to a wage of $1.50 causes projects 9, 10, and 11 to pass out of the "reject" category (given the distributional weight of 2.0). And with a further modest rise in the project wage (say to $1.70), project 12 would also easily pass the social cost-benefit test![7]

[7]Obviously, there is a risk, if distributional weights are taken at face value and are seriously and conscientiously applied, of projects being "made" acceptable simply by raising the wage or other reward paid to factors whose $\phi_i > 1$. Even if this risk is somehow surmounted, subtle pressures will exist for finding any plausible excuse to generate economic rents by paying above-market remunerations to specified factors. This causes projects to be accepted which would be rejected on standard criteria and thus represents another way in which the use of distributional weights causes income transfers to be made—but not without a cost in terms of efficiency.

V. Distributional Weights and the Income Tax

The criterion most often mentioned as the basis for assigning distributional weights is income, the presumption being that, the higher the income of an individual (or spending unit), the lower will be the distributional weight (ϕ_i) to be applied. This naturally brings to mind the income tax, as a possible way of directly—once and for all, as it were—taking the weights into account, rendering it unnecessary thereafter to bring them into play when analyzing commodity taxes, individual projects, etc. Unfortunately, the world is not so simple.

A. Consider first the case where the supply of labor is zero elastic with respect to the wage and that of saving with respect to the rate of return thereon. These assumptions rule out marginal adjustments of either leisure or savings and in effect turn the income tax into something like a variable head tax or capital levy, with no incentive effects or efficiency costs whatsoever.

This would seem to be fertile soil for a progressive income tax, and many economists in fact believe that such a tax is what would emerge (analytically) from an application of distributional weights analysis in this case. Alas, as many other economists remember from their reading of Pigou and other authors, this is not true. Declining marginal social utility of income leads (in the absence of incentive effects) not to progressive but to confiscatory taxation, so as to bring about the full and complete equalization of income. If there were no costs associated with the tax-transfer mechanism, the optimal policy under distributional weights would be to tax away all income in excess of $\bar{Y}$ (the average over all the relevant units) and use the proceeds to bring every unit whose income is below $\bar{Y}$ up to that level. If the tax-transfer mechanism carries with it some administrative costs, the optimal policy would presumably be to tax away all income in excess of $\bar{Y} + \delta$ and supplement lower incomes up to $\bar{Y} - \delta'$, where δ and δ' play the same role as transport costs in international trade theory. Basically, the end result of income equalization remains, only slightly modified.

B. Now consider the case where the supply of labor of any given type is upward sloping but where each type of labor is a noncompeting group vis-à-vis the other types. That is to say, the upward slope of the supply curve of each type of labor reflects substitution between labor and leisure of that particular group, not shifts of labor from one group (type, class) to another.

Assume, too, that the labor types are arrayed in a determinate order with respect to the incomes they earn, type 2 being richer than type 1, etc., and type N being the richest. To each group we can therefore assign a distributional weight, with $\phi_1 > \phi_2 > \cdots > \phi_{N-1} > \phi_N$.

For the moment let us rule out the idea of taxing (for distributive reasons) persons with distributional weights greater than one; our analysis will therefore consider (initially) only those groups having distributional weights less than one. Group 1 has a weight ϕ_1 only moderately less than 1; group N has the lowest weight.

If we set ourselves the problem of determining the "proper" flat-rate tax to apply to each of these noncompeting groups, the analysis of figure 2*a* holds precisely, assuming that the demanders of each group's services have "average" distributional weights ($\phi_d = 1$). A rise in the rate of income tax (t_i) in group i produces an uncompensated loss to the government ($G-$) equal to $t_i w_i^g \, dL_i$, while engendering transfers to the government of $L_i dw_i^g$ from demanders, and of $L_i \, dw_i^n$ from suppliers, of L_i, respectively.[8] The loss to demanders nets out against the corresponding gain to the government, but the transfer between suppliers of L_i and the government occasions a social gain of $(1-\phi_i)L_i dw_i^n$. The optimum tax on group i is reached when $(1-\phi_i)L_i dw_i^n = w^g \hat{t}_i dL_i$, or where, defining ε_i as the elasticity of labor supply for group i, $\hat{t}_i = (1-\phi_i)(w^n/w^g)/\varepsilon_i$.[9] Given that $w^n = w^g(1-t_i)$, this reduces to $\hat{t}_i/(1-\hat{t}_i) = (1-\phi_i)/\varepsilon_i$.

If the elasticity of labor supply were the same for all groups, the resulting pattern of taxation would in a sense be progressive, for the higher rates ($\hat{t}_i$) would apply to the groups with the lower weights (ϕ_i). But the result would not be a true income tax in the usual sense, since there would be no common schedule applicable to all income earners. Each group would face a proportional tax at "its" own rate, and those rates would turn out (with all $\varepsilon_i = \bar{\varepsilon}$) to be higher for the higher-income groups.

C. The story is somewhat different for a "true" income tax—that is, one in which a single schedule applies to all groups. Here we consider a schedule which consists of linear segments, t'_1 applying as the marginal rate in the income bracket immediately above the minimum taxable income, t'_2 as the marginal rate in the next higher bracket, etc. To determine the optimum tax for bracket i (associated with income from L_i), we still have an uncompensated cost to the government equal to $t'_i w_i^g dL_i$. But the associated transfer is now not $L_i dw^n$ but, rather, $1/2 \; N_i B_i \, dt'_i + \sum_{j>i} N_j B_i \, dt'_i$ (where B_i is the width of the ith bracket and N_j is the number of taxpaying units in the jth bracket). This simply says that, when the tax applying between \$1,000 and \$2,000 is raised from 11 percent to 13 percent, everybody with incomes above \$1,000 pays more tax. Each of those with incomes above \$2,000 pays exactly \$20 more, while each of

[8] Here w_i^n is the net-of-tax wage, w_i^g the gross-of-tax wage, and L_i the amount of labor of type i transacted in the marketplace.

[9] The fact that it is $\hat{t}_i/(1-\hat{t}_i)$, not the tax rate itself, that is equal to $(1-\phi)/\varepsilon_i$ is due simply to the fact that the tax rate is here defined as applying to the gross (rather than the net) wage.

those with incomes between \$1,000 and \$2,000 pays 20 percent of the excess over \$1,000.[10]

The benefit associated with these transfers is divided as follows:

1. A part ($=\frac{1}{2}N_i B_i\, dw_i^q$) is a transfer between demanders of type i labor and the government. Since the demanders (direct and indirect) of each type of labor are here assumed to be a representative sample of the whole population, having a distributional weight of 1 (equal to the government's), no distributional benefit or cost is associated with this transfer.

2. A part ($=-\frac{1}{2}N_i B_i\, dw_i^n$) is a transfer from suppliers of type i labor to the government.[11] The social benefit of this transfer is calculated by applying to it a weight of $(1 - \phi_i)$.

3. There remain a series of other transfers, $N_j B_i\, dt_i'$, to the government from taxpayers in the income brackets $(j > i)$ above the one for which the marginal rate change is being assessed. To each of these transfers, the weight of $(1 - \phi_j)$ is applied.

Table 1 carries out the necessary calculations for a simple example, supposed to represent a hypothetical LDC with an income tax that strikes only in the upper part of the income distribution. The lowest bracket covered by the income tax is that from \$3,000 to \$5,000. The distributional weight applying to this group is 0.9; therefore, a transfer of 1 from this group to the government generates a "distributional benefit" of 0.1. For higher income brackets, the relevant distributional weights decline, until finally, for the group with incomes between \$18,000 and \$32,000, a transfer of 1 to the government generates a "distributional benefit" of 0.5.

The table seeks to elucidate the calculation of the optimal income tax in this situation. The hypothetical experiment is always a rise of one percentage point in the marginal tax rate, t_i', applying to a specific group. The assumption made regarding the reaction of labor supply to this disturbance is simply that the equilibrium level of L_i falls by 0.5 percent.[12]

[10]The factor 1/2 applied to $N_i B_i dt_i'$ makes the approximation that the average taxable income in the ith bracket is precisely at the midpoint of the bracket. Needless to say, this is only an approximation, and the formula could easily be adjusted to fit the facts of any given case, with any desired degree of precision.

[11]Recall that dw_i^n is negative for an increase in t_i.

[12]Elasticities of demand and supply that are approximately equal come close to yielding this result. The precise result is obtained when $[(w_i^n/w_i^q) + (\varepsilon_i/|\eta_i|)] = 2$ (ε_i is the elasticity of supply, η_i the elasticity of demand [< 0] for labor of type i).

TABLE 1

DETERMINATION OF OPTIMAL INCOME TAX FOR GROUPS WITH DISTRIBUTIONAL WEIGHTS LESS THAN UNITY

A. "INFRAMARGINAL" BENEFITS DUE TO TRANSFERS FROM HIGHER INCOME GROUPS

Income Bracket ($)	N Units (1)	Distributional Weight (ϕ_i) (2)	Distributional Weight ($1-\phi_i$) (3)	Distributional Benefit ($) from Transfer of 20 (4)	30 (5)	40 (6)	60 (7)
3,000–5,000	80	.9	.1	...	...	...	...
5,000–8,000	60	.8	.2	240	...	...	...
8,000–12,000	40	.7	.3	240	360	...	...
12,000–18,000	25	.6	.4	200	300	400	...
18,000–32,000	15	.5	.5	150	225	300	450
Total	220	...	...	830	885	700	450

B. COSTS AND BENEFITS ASSOCIATED WITH OWN GROUP

Income Bracket ($)	N Units (1)	Total Income in Group ($) (2)	Social Cost Due to 1/2% Reduction in Labor Supply ($) (3)	Social Benefit Due to Transfer from Own Group: Avg. Transfer ($) (4)	($1-\phi_i$) (5)	Own Group Benefit ($) (6)
3,000–5,000	80	320,000	1,600 t_1'	10	.1	80
5,000–8,000	60	390,000	1,950 t_2'	15	.2	180
8,000–12,000	40	400,000	2,000 t_3'	20	.3	240
12,000–18,000	25	375,000	1,875 t_4'	30	.4	300
18,000–32,000	15	375,000	1,875 t_5'	70	.5	525

C. CALCULATION OF OPTIMAL TAX

Incremental Cost	=	Inframarginal Benefit	+	Own Group Benefit	=	Incremental Benefit
1,600 t_1'	=	830	+	80	=	910
1,950 t_2'	=	885	+	180	=	1065
2,000 t_3'	=	700	+	240	=	940
1,875 t_4'	=	450	+	300	=	750
1,875 t_5'	=	0	+	525	=	525

Optimal Tax Rates...............

$\hat{t}_1' = 910/1{,}600 = 57\%$
$\hat{t}_2' = 1{,}065/1{,}950 = 55\%$
$\hat{t}_3' = 940/2{,}000 = 47\%$
$\hat{t}_4' = 750/1{,}875 = 40\%$
$\hat{t}_5' = 525/1{,}875 = 28\%$

In table 1A, the income brackets, numbers, and distributional weights relevant for the various groups are presented. In the right-hand columns of this section the calculation is made of the distributional benefit associated with the "inframarginal" transfers (from higher-income groups to the government) that are occasioned by a one-percentage-point rise in each given bracket rate. For example, a rise of one percentage point in the marginal rate applying to the first bracket means that each higher income group will have to pay $20 more in tax. For the second group this transfer carries a benefit of $20 × (1 − 0.8) = $4 for each unit, or $240 for all 60 units. The same transfer of $20 carries a per unit benefit of $20 × (1 − 0.5) = $10 for the highest income group, or $150 for all 15 units in that group.

In table 1B are shown the costs and benefits associated with the particular group for which the rise in tax rate is assumed. Recall that the labor supply of any group is supposed to be reduced by 0.5 percent in response to a one-percentage-point rise in that group's marginal tax rate. We measure labor supply in "dollars worth," so that a 0.5 percent reduction in the labor supply of the first group is measured as a reduction of $1,600 (=0.005 × $320,000), while for the last group a similar 0.5 percent reduction is measured as $1,875 (= 0.005 × $375,000). The efficiency cost of a small rise in the tax is simply this induced reduction in labor supply times the marginal tax rate of the group in question; it is shown in column 3.

In columns 4–6 is measured the social benefit of the transfer from each group to the government as a consequence of a rise of one percentage point in that group's own tax rate. The average transfer given in column 4 assumes that the members of the income class are symmetrically distributed about its midpoint. Thus in the first group a taxpaying unit with $3,500 of income would pay only $5.00 of extra tax when the rate on income above $3,000 was raised by one point. However, with a symmetrical distribution, to each such family there would correspond another one with $4,500 of income, which would pay $15 of extra tax. In this way we can impute to the group as a whole an average transfer equal to 1 percent of half of the bracket width. The benefit associated with this transfer (shown in col. 6) is equal to the transfer itself (col. 4) times the benefit per dollar transferred ($1 - \phi_i$, given in col. 5), times the number of taxpaying units in the group (col. 1).

Table 1C presents the calculation of the optimal marginal tax rate for each group. This is done by setting the incremental cost of a one-point rise in the marginal rate against the sum of the two types of incremental benefit—the inframarginal benefit associated with the transfers from higher-income groups (table 1A, totals) plus the social benefit accruing as a result of transfers from the same group for which the tax is being set (table 1B, col. 6).

The results of the calculation show optimal tax rates ranging down from 57 percent for the lowest income group shown in the table to 28 percent for the highest. That is to say, in spite of distributional weights that decline from 0.9 to 0.5 as income rises, the optimal schedule of marginal tax rates for the income groups shown is uninterruptedly regressive! The reason for this result is not hard to find—it lies in the "triangular" pattern of columns 4–7 of table 1A. When the marginal tax rate is raised for a given income group, all higher groups also pay the extra tax on that particular bracket. When the given group is low on the income scale, there are many higher groups that pay this extra tax. When a high-income group is concerned, however, the number of groups still higher on the income scale will certainly be fewer, and may be (in the case of a tax rise for the highest bracket) none at all.

D. The example of Section V.C above considered the setting of an optimal tax schedule for groups with distributional weights less than unity. Though the restriction of income taxation to such groups may seem at first glance so obvious as to go without saying, further reflection reveals that no income-tax system based on the strict application of distributional weights would meet this criterion. Every income-tax system that is fully consonant with distributional weights that decline with income must assess positive marginal tax rates against at least some income ranges for which the distributional weight is greater than or equal to unity.

This proposition follows from precisely the same characteristic that tends to make the optimal tax system regressive among groups with distributional weights that are less than 1. Consider the income group whose distributional weight is precisely equal to 1. There is (by definition) no "own group" benefit stemming from the transfer of tax funds between this group and the government. There is, however, an "inframarginal" benefit from raising the marginal tax rate above zero for this group—for such a rise will cause the higher-income groups to pay more tax, without at the same time distorting their marginal choices between effort and leisure. Since this inframarginal benefit, from a small rise in the group's tax rate above zero, is finite (or of the first order of smalls), while the social cost due to the reduction in the group's labor supply is infinitesimal (or of the second order of smalls), it will always pay society to set the tax on this group above zero.

With respect to groups with weights that are strictly greater than 1, similar considerations apply, modified by the facts that the own group benefit is now necessarily negative and that the inframarginal benefit will now normally have some negative components.[13] Nonetheless, barring quite unusual discontinuities either in the pattern of distributional weights or in the income distribution itself, there will typically be a substantial range of below-average incomes which "qualify" for significant and positive marginal tax rates.

Such a case is illustrated in table 2, which can be regarded as a continuation of table 1 to cover groups with distributional weights greater

[13]For the first group with distributional weight (just barely) greater than one, all higher-income groups will have weights less than or equal to one. In this case the own group benefit will be negative, but all components of the inframarginal benefit will be greater than or equal to zero.

TABLE 2

DETERMINATION OF OPTIMAL INCOME TAX FOR GROUPS WITH DISTRIBUTIONAL WEIGHTS ≥ 1

A. "INFRAMARGINAL" BENEFITS DUE TO TRANSFERS FROM HIGHER INCOME GROUPS

Income Bracket ($)	N Units (1)	Distributional Weight (ϕ_i) (2)	Distributional Weight $(1-\phi_i)$ (3)	Distributional Benefit ($) from Transfer of 10 (4)	4 (5)	6 (6)	10 (7)
0–1,000	120	1.4	−.4	...	...	...	...
1,000–1,400	60	1.2	−.2	−120	...	...	...
1,400–2,000	80	1.1	−.1	−80	−32	...	...
2,000–3,000	100	1.0	.0	0	0	0	...
3,000–5,000	80	.9	.1	80	32	48	80
5,000–8,000	60	.8	.2	120	48	72	120
8,000–12,000	40	.7	.3	120	48	72	120
12,000–18,000	25	.6	.4	100	40	60	100
18,000–32,000	15	.5	.5	75	30	45	75
Total	580	...	...	295	166	297	495

B. COSTS AND BENEFITS ASSOCIATED WITH OWN GROUP

Income Bracket ($)	N Units (1)	Total Income in Group ($) (2)	Social Cost Due to 1/2% Reduction in Labor Supply ($) (3)	Social Benefit Due to Transfer from Own Group: Avg. Transfer ($) (4)	$(1-\phi_i)$ (5)	Own Group Benefit ($) (6)
2,000–3,000	100	250,000	1,250 t_0'	5	.0	0
1,400–2,000	80	136,000	680 t_{-1}'	3	−.1	−24
1,000–1,400	60	72,000	360 t_{-2}'	2	−.2	−24
0–1,000	120	60,000	300 t_{-4}'	5	−.4	−240

C. CALCULATION OF OPTIMAL TAX

Incremental Cost	=	Inframarginal Benefit	+	Own Group Benefit	=	Incremental Benefit
1,250 t_0'	=	495	+	0	=	495
680 t_{-1}'	=	297	+	(−24)	=	273
360 t_{-2}'	=	166	+	(−24)	=	142
300 t_{-4}'	=	295	+	(−240)	=	55

Optimal Tax Rates ....

$\hat{t}_0' = 495/1{,}250 = 40\%$

$\hat{t}_{-1}' = 273/680 = 40\%$

$\hat{t}_{-2}' = 142/360 = 40\%$

$\hat{t}_{-4}' = 55/300 = 18\%$

than 1. (For convenience, the subscripts used to distinguish the different income groups are keyed to their distributional weights, so that the group with a distributional weight of 1 is given a subscript of zero; that with a distributional weight of 1.1 has a subscript of −1; that with a distributional weight of 1.4 has a subscript of −4, etc.)

Table 2A shows the inframarginal benefits generated by a one-percentage-point increase in the tax on the lower-income groups. Column 4 shows the effects of a transfer of $10 per taxpaying unit, a rising of a one-percentage-point increase in the rate in the $0–$1,000 income group. In column 5 are displayed the effects of a transfer of $4.00 per taxpaying unit, generated by a one-percentage-point rise in the marginal tax rate of the $1,000–$1,400 income group. Similar interpretations should be given for the figures in columns 6 and 7, here referring to rises in the tax rates on the $1,400–$2,000 and the $2,000–$3,000 income groups, respectively.

Table 2B follows directly the format of the corresponding part of table 1, using (for col. 3) some assumption concerning labor supply—namely, that for each group a one-percentage-point rise in its tax rate causes a fall of 0.5 percent in its labor supply. The own group benefit shown in column 6 is naturally zero for the group whose distributional weight is unity and negative for those with distributional weights of less than 1.

Table 2C displays the calculation of the optimal marginal tax rate for each group. In spite of the fact that the direct distributional benefit of transfers from groups 0, -1, and -2 to the government is zero or negative, it is surprising to find that the optimal marginal tax rate for these groups is 40 percent—equal to that (t'_4) corresponding to the second highest income bracket ($12,000–$18,000) and higher than that (t'_5) corresponding to the highest income group (see table 1C).

Thus we can see quite clearly that the use of distributional weights that vary inversely with income need not give rise to a generally progressive tax structure and can very well conduce to relatively high income taxes even for the poor.[14]

It should be noted, too, that the weights employed in tables 1 and 2 had a relatively modest range of variation and that the general attributes of the hypothetical economy there treated were not unrealistic, at least for the poorer countries of the world.[15]

VI. What Is Wrong with Distributional Weights?

There can be no doubt that distributional weights (particularly weights which vary inversely with income) have a strong appeal to those nurtured in the grand tradition of economics. The notion of a diminishing marginal utility of income pervades the classical and neoclassical economic literature; and although most economists have come to shun interpersonal

[14] The exploration here did not discuss the possibility of sharp jumps in the function relating income tax payments to income.

[15] The top 2½ percent of workers in this example get about 15 percent of the income; the top 7 percent get about 30 percent of the income. In contrast, the bottom 20 percent of the units get about 2½ percent of the income and the bottom half about 15 percent of total income. The Gini coefficient is .4, not unusual among the less-developed countries.

comparisons of utility whenever they feel they can avoid or circumvent them, they usually make them when they have to. And when they do, they most commonly rely on the assumption that the loss of a dollar means less to the rich man than to his poorer brother.

Now, it is difficult indeed to tell the difference between a system of distributional weights that decline with income and an alternative system of measurable utility made interpersonally comparable on some basis and characterized by diminishing marginal utility of wealth or income. In short, I think it is fair to say that the notion of distributional weights (that decline with income) has been lurking for a long time in our professional unconscious.

Yet the implications for policy of a thorough and consistent use of distributional weights turn out to be quite disturbing—even to those who (perhaps as today's representatives of our utilitarian heritage) are attracted to the idea. These implications include

1. With respect to commodity taxation (see Secs. II, III.A, III.B), it is only by accident that an optimal excise tax (or subsidy) would be zero. For most commodities, the distributional effects of a marginal increment of tax would be different from zero, starting from an untaxed, unsubsidized situation. An optimum would be found only when the point was reached where the distributional gains from a small further increase in the tax or subsidy in question were just matched by the additional efficiency losses it entailed.

2. With respect to investment projects (Sec. IV), the implications of distributional weights are quite similar; where the estimated net effect of a project is a transfer to people with distributional weights higher than 1, the distributional benefit assigned on this ground represents the price that society "should" (by the logic of distributional weights) be willing to pay in terms of economic efficiency and still approve the project. That is to say, when a project carries distributional benefits, the straightforward application of distributional weighting schemes would lead us to use as a criterion the balancing (at the margin) of distributional benefits against efficiency costs.

3. Finally, with respect to income taxation (Sec. V), we find that such taxation can readily turn out to be regressive over significant ranges of income, that it will always apply positive marginal tax rates to groups from which the direct distributional consequence of a transfer to the government is

> adjudged to be zero, and that it will quite commonly entail significant marginal tax rates at income levels well below the average. All of these consequences stem from the fact that a rise in the marginal income-tax rate for a lower-income group results in higher inframarginal taxes for higher-income groups.
>
> The distributional benefit of a rise of one percentage point in the marginal tax rate is greatest at that point in the income distribution where the distributional weight applying to the taxpayers and that applying to the government are just equal, for this rise will draw additional taxes from all groups with higher weights. The distributional benefit of a rise in the highest-bracket rate stems only from the highest-income group itself; it carries no inframarginal benefit from higher brackets. Small wonder, then, that regressive taxation can emerge from a straightforward application of distributional weights.

Now, these implications of distributional weights need not in principle be very disturbing. If the differences in the weights assigned to different groups were small enough, the implied extra burdens in terms of efficiency would also be small. But when people talk seriously about making welfare economic analysis more meaningful through the use of distributional weights, they simply do not have in mind weights of 1.1 for the pauper and 0.9 for the millionaire.

In fact, applications in which specific forms of the distributional weighting function are used (usually by way of examples) most commonly employ a logarithmic functional form, such as $\phi_i = (y_i/\bar{y})^{-\lambda}$, where y stands for income, or alternatively, $\phi_i = (c_i/\bar{c})^{-\lambda}$, where c stands for consumption. These functions have the attribute that the difference between the weights attaching to different individuals can be exceedingly large (in mathematical terms it is unbounded).

Numerical examples of how the weights can vary are given in Little and Mirrlees (1974). For an income (or consumption level) that is one-fourth of the "norm" (for which ϕ_i is equal to 1), the weight is 4 if λ is 1, 16 if λ is 2, and 64 if λ is 3. In the latter case a transfer from a typical family to one that was only one-fourth as well off could entail a waste of up to 63 times the amount transferred and still be, in principle, acceptable. For $\lambda = 1$, such a transfer would still be acceptable even if it entailed a waste up to three times the amount transferred.

The above estimates of "acceptable" levels of waste under distributional weights assume that the transfer is fully effectuated—that is, that the beneficiaries get what the transferors lose, with the "waste" taking place elsewhere in the economy's machinery. A more conservative assumption is that the "waste" takes place in the process of effectuating the transfer itself, with the beneficiaries actually receiving only a fraction of what the transferors give up. This comes closer to what actually happens with inefficient investment projects or other economic policies (such as taxes) or programs (such as the U.S. sugar program).

An example I have found useful for conveying the essential elements of this sort of transfer is a project to send ice cream on camelback across the desert from a richer oasis to a poorer one. If λ is equal to 1 and the richer oasis has twice the per capita income or consumption of the country as a whole (this makes its $\phi_i = 1/2$), while the poorer oasis has a per capita income equal to half the national average (for a ϕ_i of 2), it would be possible for up to three-fourths of the ice cream to melt en route without causing the project to fail the distributionally weighted cost-benefit test. (The resource costs of the camel transport are neglected in this example.) If λ were equal to 3, the ϕ_i of the richer oasis would be 1/8 and that of the poorer one would be 8. In this case (again neglecting the resource costs of transport), up to 63/64 of the ice cream could melt away without causing the project to fail the test.

The lesson from these examples is clear: When distributional weights are used together with weighting functions of the type most commonly employed in writings on the subject, the result is to open the door to projects and programs whose degree of inefficiency by more traditional (unweighted) cost-benefit measures would (I feel confident) be unacceptable to the vast majority of economists and of the informed public.

The dilemma, then, is that, when the differences in weights are small, distributional considerations are reflected only to a minor degree in the evaluation process. When, on the other hand, the differences in weights get to be large, it is all too easy for considerations of distribution to swamp those of efficiency altogether, and for grossly inefficient policies, programs, and projects to be deemed acceptable.

VII. In Search of a Solution

There are a variety of ways in which one can approach the dilemma presented in the preceding section (and, indeed, throughout this paper). I shall here present three distinct alternatives, the main differences among them being their degree of fidelity to the basic idea that the right way to take distributional considerations into account in social cost-benefit analysis is by way of distributional weights.

A. Imposing Both Tests Simultaneously

One sure way to avoid paying an exaggerated price in terms of lost efficiency for the redistributive benefits that a weighting scheme would bring into account is simply to require that policies and projects should pass both tests—the pure efficiency test imposed by traditional applied welfare economics plus the weighted test that emerges when the welfare gains and losses of different groups are multiplied by designated weights before the balance is struck.

Of course, the requirement that both tests be passed does not mean that the distributional gains that result from applying the dual test come as a "free good." As in most other areas of economics, a price must be paid for the benefit obtained.

Consider the classic problem of minimizing the efficiency cost of raising a specified sum of money via excise taxes on a subset of commodities (Ramsey 1927). Its counterpart with the full use of distributional weights would entail maximizing the weighted sum of the surpluses of the different groups or categories that were relevant for the problem (see Diamond 1975). This "unconstrained" application of distributional weights might have a very high efficiency cost, measured in the traditional way. In contrast, the simultaneous application of the two sets of criteria might lead us to minimize the traditional efficiency cost, subject to the constraint that a weighted sum of the net benefits accruing to different individuals and groups be greater than zero (or some other number). If the unconstrained cost minimization meets the constraint imposed, the result of applying the dual criterion is the same as that for the standard approach. If, however, the constraint is not met under a straightforward minimization of efficiency cost, then efficiency cost in the constrained case will be greater. Its excess over the cost in the unconstrained case is the price paid for the improved distributional characteristics of the constrained solution.

If both tests were imposed simultaneously in a project-evaluation setting, the result would be similar. If the critical rate of discount would be 10 percent in an unconstrained application of traditional cost-benefit analysis, it might here turn out to be 7 percent or 8 percent. Alternatively, if a given rate of discount were used throughout, and a criterion ratio of discounted benefits to discounted costs imposed (as is traditionally done in cases of limited capital budgets), then the imposition of a distributional constraint might cause this critical ratio to fall, say, from 1.40 to 1.25 for benefits and costs measured in the traditional, unweighted way.

It can be seen, then, that the simultaneous application of both weighted and unweighted tests is not costless. Nonetheless, there is a sense in which this method minimizes the costs entailed in bringing distributional weights into the picture. That, indeed, is its chief virtue.

If the dual criterion is adopted, special attention should be paid to the way in which so-called separable components of programs or projects are handled. A very strict application of the dual criterion might insist that each separable component pass both unweighted and weighted cost-benefit tests. Less stringent would be a rule insisting that separable components should continue to pass, as in the traditional approach, the unweighted cost-benefit test, while requiring a nonnegative weighted net benefit only for the project as a whole. In this case individual components could have weighted costs greater than weighted benefits, though they would not be permitted to have unweighted costs greater than unweighted benefits. This type of hybrid criterion implicitly gives priority to the efficiency criterion; the opposite hybrid—insisting that each separable component pass the weighted test while only the global program or project must pass the unweighted test—gives relative priority to the distribution criterion.

B. A Single Premium Magnifying the Net Benefits (Costs) of Those Below a Poverty Line

The path that leads to this "solution" is devious but (I think) interesting. If one wants to justify the traditional cost-benefit procedures (which neglect distributional considerations), the best approach is surely to adopt the assumption that lump-sum transfers can be effectuated at will and without any cost in terms of economic efficiency. This is, at least in one dimension, the key assumption underlying the Hicks-Kaldor principle under which programs or projects are to be approved as long as the gainers could in principle compensate the losers and still be no worse off than before.

Of course, such compensations and transfers are rarely made—a fact that is subject to at least two interpretations: (*a*) "society" could costlessly effectuate the transfers but chooses not to do so because it is satisfied with the status quo, and (*b*) the compensations and transfers are not costless after all. Clearly, the second of these reasons carries more weight (and with more people) than the first.

Yet the fact that it costs more than zero to effectuate a transfer—once a decision in that direction has been reached—does not mean that we need to accept resource waste of one-half or three-fourths or seven-eighths or 63/64 of the amounts involved in order to bring the transfer about. While the most eminent minds in the economics profession would be hard put to find ways of effectuating quantitatively important transfers at efficiency costs of less than 1 percent of the amounts involved, most beginning graduate students could in one afternoon invent a hundred ways of bringing about major transfers at an efficiency cost of less than 20 percent of the amounts involved.

On this reasoning, it seems quite sensible to think of there being something like a critical level (say, 10 percent) of efficiency cost per dollar transferred. If a way is found of effectuating desirable transfers more cheaply than this—either through a conscious program or as a side effect of other programs or projects whose main aims may be quite different—it should (on this line of reasoning) be welcomed. But no merit should accrue to a program or project for its distributive achievements if their implicit cost is above the critical level (here 10 percent of the amount transferred).

This view of the world fits neatly into modern economies. The 10 percent premium may be thought of as reflecting a shadow price of public funds that is different from unity. Alternatively, one can reach the same conclusion by following one of the cardinal principles of cost-benefit analysis: Never attribute to an action a benefit that exceeds the alternative cost of achieving the same result.

So far, this line of argument would lead us to try to make explicit the various transfers that might be implicit in a program or project or other policy decision and to array them in descending order of the net benefit that would be assigned to them using distributional weights. Where that net benefit was less than 10 percent of the amount transferred, it would stand unaltered in the cost-benefit calculation. But where the net benefit from an implicit transfer was greater than 10 percent of the amount transferred, it would be stricken from the balance and replaced by the alternative cost figure of, in the present example, 10 percent of the amount transferred.

Up to this point, we are still being quite faithful to the use of distributional weights. All we have done is to recognize the alternative cost principle and bring it to bear upon our problem. But here the argument starts to follow a different and less strictly theoretical track. Relative magnitudes become important here. How rich are the payers relative to the recipients of the implicit transfers in question? How rapidly do the distributional weights fall as income (or consumption) rises over this range? How large is the critical premium (10 percent in our example) relative to the range within which the distributional weights vary? And finally, does "society" really "care" about transfers taking place among the wealthy and the reasonably well-to-do?

The specific "facts" that we presume to exist, in making the next leap in the argument, are the following:

1. The distribution of income has a wide span—say, a factor of 10 or more.

2. Distributional weights are not all tightly clustered about 1. For a 10-fold range of income distribution, the range of weights might be 0.50–2 or 0.33–3 or 0.25–5—but *not* 0.95–1.05.

3. The critical premium is not much more than 10 percent (say, 20 percent at the outside).

4. "Society" is substantially indifferent concerning transfers among those with above-average incomes.

Presumption (4) enables us to neglect transfers within the well-to-do group. Presumptions (1) and (2) indicate that transfers from those significantly above a "poverty line" to those below it will carry substantial benefits when distributional weights are used. Presumption (3) says, however, that these benefits will be supplanted, in the cost-benefit calculus, by the standard premium (10 percent?) based on the alternative cost of effectuating transfers.

This critical premium dominates the likely calculus of distributional costs and benefits to such an extent that the precise pattern of decline of the distributional weights as income, consumption, or wealth rise turns out to be quite unimportant. In other words, it does not matter much whether λ, as defined in Section VI, has a value of 0.5 or 1 or 3 or 5—a 10 percent premium based on the alternative cost of effectuating transfers would likely dominate the cost-benefit calculus in any event.

All of this leads to a robust and (to me at least) appealing simplification—to work with standard (unweighted) cost-benefit criteria in general, but to calculate also a net change (plus or minus) in surplus for the group below the poverty line. The alternative cost premium times this net change would then be treated as the net distributional benefit, or cost, of the program under study.

This simplification has the advantage of reducing the estimation of distributional effects to something much less than (and therefore much more feasible of estimation than) a full catalog of how every group's surplus would be affected. It also has the advantage of being quite adaptable to different ideologies or visions of the good society. A conservative with a strong belief in the work ethic might set the poverty line in the United States at $2,000 and the alternative-cost premium at 3 percent. A soft-hearted liberal, on the other hand, might set the poverty line at $10,000 and the alternative-cost premium at 20 percent. Between these two there is ample range to capture the implicit value systems of, say, Goldwater or Reagan on the one hand as well as Humphrey or McGovern on the other. From my own discussions with professional colleagues on these matters, I find that the more seriously they contemplate the full range of issues and problems involved, the more they incline toward this solution as against that of Subsection A.

C. New Light on the Traditional Solution

Having come this far, we may better appreciate some of the pitfalls that await efforts to take systematic account, within the formal procedures of social cost-benefit analysis, of distributional and other nonefficiency considerations. In this concluding subsection, I shall only sketch a line of thought that can quite readily lead to a reaffirmation of the traditional efficiency-oriented calculus.

We begin with a hypothetical experiment, or, alternatively, an appeal to our own experience and observation, concerning transfers within family groups—or, even better, within groups of grown siblings. What sort of circumstances elicit major transfers? How are these transfers effectuated? Could they be rationalized or justified in terms of a set of distributional weights applied by and within the family?

Taking the last question first, let me emphatically deny that the behavior of siblings can be plausibly rationalized or explained in terms of distributional weighting schemes of the kind we have been analyzing. Such schemes would in principle call for all windfalls to be shared among siblings—rises in pay, stock market successes, lottery winnings, etc. Yet we do not observe systematic transfers taking place in such cases.

The transfers that we do observe among siblings have much more to do with major negative windfalls than with positive ones. The family coalesces to help when a brother loses his job or a sister has a serious accident or a house burns down, etc.—but not when an expected raise fails to come through or when a sibling's automobile is smashed up.

Looking deeper into transfers between siblings, they seem to be oriented more toward financing specific expenditures or "needs" rather than toward providing a generalized sharing of a common budgetary fund. Basic food, clothing, and shelter in a time of unemployment or distress, emergency medical care at almost any time, education of the family's children up to some accepted level—these appear to be the needs that call forth intrafamily transfers with the greatest reliability.

Such transfers, then, are typically not generalized and systematic "sharing of the wealth" transfers in cash, such as would be implied by commonly suggested distributional weighting functions. Instead, they tend to be transfers that are, for all practical purposes, "in kind" and oriented toward meeting a certain minimum (which obviously varies from case to case) standard of food, clothing, shelter, medical care, and education.

Having drawn this conclusion from a hypothetical look at family behavior, I for one was quite surprised at how well it fit the observed pattern of societal transfers. Free public education, minimum medical care for those who cannot afford it, food and minimal shelter for the destitute—these all have roots going back long before distributional weights were invented.

In contrast, the economist's favorite gimmick—the cash transfer—has had rough going in virtually every modern society. Free education, subsidized medical care, food and housing for the poor—these are all standard fixtures in today's world—developed and developing countries alike. Efforts to convert these subsidies to a cash equivalent independent of the purpose for which it might be spent have everywhere faltered and failed. Might it not be that there is a message here for economists trying to grapple with distributional questions?

I believe so, and the message really tells us to be more modest and humble, less ready to arrogate to our profession and discipline the solution of all of society's problems.

There are plenty of grounds for modesty and humility as we approach these basic issues. We may all agree that the poor should be helped, but who are the poor? Is income a test? Most graduate students are poor by that standard. We want to help alleviate the human suffering connected with unemployment without inducing people to go casually and habitually "on the dole"—as many do. We want to see to it that the needy have an adequate standard of medical care, but who is to say that the long waiting lines observed in the public clinics of many countries do not perform a quite useful function in restricting the beneficiaries of the free service to people who are really poor and really ill?

All that need be said here is that distributional weights, as commonly understood in our professional literature, simply do not enter into these dimensions. And the more important or relevant are these dimensions of the problem, the less useful are distributional weights as a way of grappling with it.

In the end, then, we cannot condemn as crass or unfeeling the idea our profession's possibly moving toward a "consensus" based on the traditional criterion of efficiency. On the contrary, such a result might well reflect a greater and more sensitive understanding of the value systems of our citizens and our societies, as well as a more modest and realistic appreciation of our own professional role.

Bibliography

Limitations of space as well as considerations of the scope and purpose of the present paper have precluded an adequate review of the growing literature on distributional weights. In recognition of this fact, the following partial bibliography is provided.

Atkinson, Anthony B. "On the Measurement of Inequality." *J. Econ. Theory* 2 (September 1970): 244–63.

Atkinson, Anthony B., and Stiglitz, J. E. "The Structure of Indirect Taxation and Economic Efficiency." *J. Public Econ.* 1 (1972): 97–119.

———. "The Design of Tax Structure: Direct versus Indirect Taxation." Technical Report no. 199, Stanford Univ. Inst. Math. Studies Soc. Sci., February 1976.

———. *Lectures on Public Economics*. Forthcoming.

Boskin, Michael J. "Optimal Tax Treatment of the Family." Memorandum no. 143, Stanford Univ. Res. Center Econ. Growth, 1973.

Diamond, Peter A. "A Many-Person Ramsey Tax Rule." *J. Public Econ.* 4 (August 1975): 335–42.

Diamond, Peter A., and Mirrlees, J. A. "Optimal Taxation and Public Production: Tax Rules. I." *A.E.R.* 61(March 1971): 8–27. (*a*)

———. "Optimal Taxation and Public Production: Tax Rules. II." *A.E.R.* 61 (June 1971): 261–78. (*b*)

Feldstein, Martin S. "Distributional Equity and the Optimal Structure of Public Prices." *A.E.R.* 62 (March 1972): 32–36. (*a*)

———. "Equity and Efficiency in Public Pricing." *Q.J.E.* 86, no. 2 (1972): 175–87. (*b*)

Harberger, Arnold C. "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay." *J. Econ. Literature* 9 (September 1971): 785–97.

Heller, Walter P., and Shell, Karl. "On Optimal Taxation with Costly Administration." *A.E.R.* 64 (May 1974): 338–45.

Little, I. M. D., and Mirrlees, J. A. *Project Appraisal and Planning for Developing Countries.* New York: Basic Books, 1974. (Esp. pp. 238–44.)

Meade, James E. *Trade and Welfare: Mathematical Supplement.* London: Oxford Univ. Press, 1955. (Esp. pp. 34–46.)

Mirrlees, J. A. "An Exploration in the Theory of Optimum Income Taxation." *Rev. Econ. Studies* 38 (April 1971): 175–208.

———. "Optimal Commodity Taxation in a Two-Class Economy." *J. Public Econ.* 4 (February 1975): 27–37.

Ramsey, Frank P. "A Contribution to the Theory of Taxation." *Econ. J.* 37 (March 1927): 47–61.

Squire, Lyn, and van der Tak, Herman G. *Economic Analysis of Projects.* Baltimore and London: Johns Hopkins, 1975. (Esp. pp. 101–17, 136–42.)

Stiglitz, J. E. "Simple Formulae for the Measurement of Inequality and the Optimal Linear Income Tax." Technical Report, Stanford Univ. Inst. Math. Studies Soc. Sci. Forthcoming.

UNIDO. *Guidelines for Project Evaluation.* New York: United Nations, 1972. (Esp. chaps. 13–14.)